The individual then could query a number of GIS data sets, including property parcels and values, school locations, police and fire station locations, risk of natural disaster (flood, earthquake, etc.), political districts, and environmental hazards, and quickly produce a customized map that shows all the data requested for the area surrounding the address given.

The Internet also is affecting political processes in the United States and around the world. Political candidates are establishing websites to communicate with voters, solicit funds, and organize volunteers. Interest groups are using email and websites to organize and express their views. In some cases, groups that would be very difficult to organize through traditional means, such as scientists or engineers in different parts of the country, can be mobilized through e-mail to express their views to Congress on a timely issue. Some groups are experimenting with voting via the Internet. See sidebar, "Internet Voting."

IT and S&E

The S&E community developed IT, in many cases for S&E applications. Scientists and engineers have been among the earliest and most intensive users of many IT applications. It is not surprising that IT has played a major role in the practice of S&E and in the evolution of S&E institutions.

Advances in computing, information storage, software, and networking are all leading to new tools for S&E, ranging from automated scientific instruments to supercomputers for modeling and simulation. IT has made possible new collections of data and new ways to access scientific information. As IT has advanced, applications for S&E have become more powerful and less expensive, and many applications, such as modeling and databases, have migrated from large mainframe computers and supercomputers to desktop computers. IT also has made possible new modes of communication among scientists, allowing them to collaborate more easily. IT affects how research

Internet Voting

Many people have expressed a strong interest in using the Internet to make voting more convenient. It is hoped that such a practice would increase participation in elections. Internet voting is seen as a logical extension of Internet applications in commerce and government. Election systems, however, must meet high standards of security, secrecy, equity, and many other criteria. These requirements make the development of Internet voting much more challenging than most commerce or government applications of the Internet.

The National Science Foundation supported a study and workshop to analyze the issues associated with Internet voting (Internet Policy Institute 2001). The study concluded that remote Internet voting (e.g., voting from the home or office) poses significant risks to the integrity of the voting process and should not be widely used in public elections until substantial technical and social science issues are addressed. On the other hand, it would be possible to use Internet voting systems at polling places, and such systems could offer greater convenience and efficiency than traditional voting systems. Voters could eventually choose to cast their ballots from any one of many polling places, and the tallying process would be both fast and certain. Because election officials would control both the voting platform and the physical environment, managing the security risks of such systems is feasible. Over time, it would also be possible to have Internet voting in kiosks—voting machines located away from traditional polling places—at convenient locations such as malls, libraries, or schools. Kiosk voting terminals pose more challenges than poll-site systems, but most of the challenges could, at least in principle, be resolved through extensions of current technology.

A broad range of research is needed on Internet voting systems. Research topics include the following:

- ♦ Approaches to meeting the security, secrecy, scalability, and convenience requirements of elections.
- ♦ Development of reliable poll-site and kiosk Internet voting systems that are not vulnerable to any single point of failure and cannot lose votes.
- ♦ Development of new procedures for continuous testing and certification of election systems, as well as test methods for election systems.
- Incorporation of human factors into design for electronic voting, including development of appropriate guidelines for designing human interfaces and electronic ballots and development of approaches for addressing the needs of the disabled.
- The economics of voting systems, including comparative analyses of alternative voting systems.
- ♦ The effects of Internet voting on participation in elections, both in general and with regard to various demographic groups, especially those with less access to or facility with computers.
- ◆ The effects of Internet voting on public confidence in the electoral process and on deliberative and representative democracy.
- The implications of Internet voting for political campaigns.
- ♦ Legal issues associated with and the applicability of existing statutes to Internet voting, including jurisdiction, vote fraud, liability for system failures, international law enforcement, and electioneering.

is conducted, how new products and processes are developed, and how technical information is communicated.

IT also is influencing technological innovation in society. These influences reflect not only changes in R&D processes but also changes in the market environment for innovation and the organization of innovative activities. Although some of these effects are most visible in the IT industry itself, IT also affects other industries, higher education, and the job market for scientists and engineers.

In general, relatively little scholarly research has been conducted on how IT affects S&E, and even less research has been performed on how IT affects innovation. This section highlights some of the limited work that has been done.

IT and R&D

IT has provided new tools for the simulation and modeling of complex natural, social, and engineering systems. It has enabled new methods of data collection and has made possible the creation of massive, complex, and shared data sets. It has changed the way scientific knowledge is stored and communicated. IT has facilitated the sharing of computational resources and scientific instruments among scientists and engineers in different locations and has aided communication and collaboration among large groups of researchers.

Advances in both hardware and software have supported new IT tools for R&D. Advances in software have been critical to the success of supercomputers that use thousands of microprocessors and have also enabled the analysis and visualization of complex problems. Software engineering also is enabling security technologies, distributed information management, high-confidence software systems, and numerous other areas of research that are needed in today's most advanced IT applications

The role of IT is not uniform across all areas of S&E. Some areas of research, such as high-energy physics, fluid dynamics, aeronautical engineering, and atmospheric sciences, have long relied on high-end computing. The ability to collect, manipulate, and share massive amounts of data has long been essential in areas such as astronomy and geosphere and biosphere studies (Committee on Issues in the Transborder Flow of Scientific Data 1997). More recently, IT has spread from its historical stronghold in the physical sciences to other natural sciences, engineering, social sciences, and the humanities and has become increasingly vital to sciences such as biology that historically had used IT less extensively.

Modeling and Simulation

Modeling and simulation have become powerful complements to theory and experimentation in advancing knowledge in many areas of S&E. Simulations allow researchers to run virtual experiments when actual experiments would be impractical or impossible. As computer power grows, simulations can be made more complex, and new classes of problems can be realistically simulated. Simulation is contributing to major advances in weather and climate prediction, computational biology, plasma science, high-energy physics, cosmol-

ogy, materials research, and combustion, among other areas. New visualization techniques for displaying simulation data in comprehensible formats have played an important role.

Simulation also is used extensively in industry to test the crashworthiness of cars and the flight performance of aircraft (U.S. Department of Energy (DOE)/NSF 1998) and to facilitate engineering design. Computer-aided design (CAD) programs can use CAD data to visualize, animate, simulate, validate, and assemble parts digitally. In some cases, CAD programs can allow a designer to insert digital representations of humans into virtual worlds to test for ergonomics, manufacturability, maintainability, safety, and style (Brown 1999). The goal of such an approach is to address these issues early in the design stage and reduce the need for physical mock-ups and rework. Both aircraft and automobile manufacturers use CAD approaches extensively.

Modeling and simulation capabilities continue to improve at a rapid rate. DOE's Accelerated Strategic Computing Initiative program, which uses simulation to replace nuclear tests, deployed the first trillion-operations-per-second (teraops) computer in December 1996. The program deployed a 12.3-teraops computer in June 2000 and plans to operate a 100-teraops computer (with 50 terabytes of memory and 130 petabytes of archival storage) by 2005 (National Science and Technology Council 1999; U.S. DOE 2001). Research funded by the Defense Advanced Research Projects Agency, the National Aeronautics and Space Administration (NASA), and the National Security Agency is evaluating the feasibility of constructing a computing system capable of a sustained rate of 1,015 teraops (1 petaflop).

Terascale computing is expected to have applications in genetic computing, global climate modeling, aerospace and automotive design, financial modeling, and other areas. To use data from human genome research, for example, new computational tools are needed to determine the three-dimensional atomic structure and dynamic behavior of gene products, as well as to dissect the roles of individual genes and the integrated function of thousands of genes. Modeling the folding of a protein to aid in the design of new drug therapies also takes extensive computing power (U.S. DOE/NSF 1998). Celera Geonomics Corporation (a genomics and bioinformatics company), Sandia National Laboratories, and Compaq entered into a partnership in January 2001 to develop algorithms and software for genomic and proteomic applications of supercomputers in the 100-teraops to 1-petaflop range, with the petaflop computer expected by 2010. Pattern recognition and data-mining software also are critical for deciphering genetic information (Regalado 1999).

Many scientists expect IT to revolutionize biology in the coming decades, as scientists decode genetic information and explore how it relates to the function of organisms (Varmus 1999). New areas of biology such as molecular epidemiology, functional genomics, and pharmacogenetics rely on DNA data and benefit from new, information-intensive approaches to research.

IT and Data

IT has long been important in collecting, storing, and shar-

ing scientific information. More recently, IT has enabled automated collection of data. For example, automated gene sequencers, which use robotics to process samples and computers to manage, store, and retrieve data, have made possible the rapid sequencing of the human genome, which in turn has resulted in unprecedented expansion of genomic databases (Sinclair 1999). In many scientific fields, data increasingly are collected in digital form, which facilitates analysis, storage, and dissemination. For example, seismic data used to measure earthquakes were once recorded on paper or film but now are usually recorded digitally, making it possible for scientists around the world to analyze the data quickly.

By 1985, 2,800 scientific and technical electronic databases (both bibliographic and numerical) already existed (Williams 1985). At that time, primarily information specialists accessed electronic databases, and many of the databases were available only for a fee. Over time, databases have expanded in number and size, and many are now widely accessible on the World Wide Web. See sidebar, "Examples of Shared Databases."

Electronic Scholarly Communication

Originally developed primarily as tools for scientific communication, the Internet and the World Wide Web continue to have a significant impact on scholarly communication in scientific and technical fields. An increasing amount of scholarly information is stored in electronic forms and is available through digital media.

Electronic Scholarly Communication Forms. Scholarly information can be placed on-line in several different forms, most of which are expanding rapidly. These forms may be classified as follows (drawing on Kling and McKim 1999 and 2000):

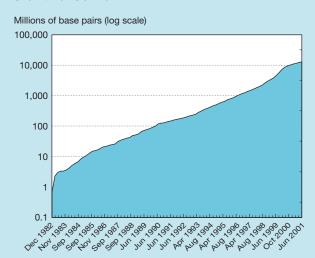
- ◆ Pure electronic journals—an edited package of articles that is distributed to most of its readers in electronic form. Examples include the World Wide Web Journal of Biology and the Journal of the Association for Information Systems.
- ♦ Hybrid paper-electronic (p-e) journals—a package of peer-reviewed articles that is distributed primarily in pa-

Examples of Shared Databases

Large shared databases have become important resources in many fields of science and social science. These databases allow researchers working on different pieces of large problems to contribute to and benefit from the work of other researchers and shared resources. Examples of such databases include the following:

- ♦ GenBank (http://www.ncbi.nlm.nih.gov/Genbank/) is the National Institute of Health's annotated collection of publicly available DNA sequences. As of June 2001, GenBank contained approximately 12.9 billion base pairs from 12.2 million sequence records. (See figure 8-22.) The number of nucleotide base pairs in its database has doubled approximately every 14 months. As part of a global collaboration, GenBank exchanges data daily with European and Japanese gene banks.
- ♦ The Protein Data Bank (http://www.rcsb.org/pdb/) is the worldwide repository for the processing and distribution of three-dimensional biological macromolecular structure data (Berman et al. 2000).
- ♦ The European Space Agency (ESA) Microgravity Database (http://www.esa.int/cgi-bin/mgdb) gives scientists access to information regarding all microgravity experiments carried out on ESA and National Aeronautics and Space Administration missions by European scientists since the 1960s.
- ♦ The Tsunami Database (<http://www.ngdc.noaa.gov/ seg/hazard/tsu.html>) provides information on tsunami events from 49 B.C. to the present in the Mediterranean and Caribbean Seas and the Atlantic, Indian, and

Figure 8-22. **Growth of GenBank**



SOURCE: Genetic Sequence Data Bank, NCBI-GenBank Flat File. Available at <ftp://ncbi.nlm.nih.gov/genbank/gbrel.txt>.

Science & Engineering Indicators – 2002

Pacific Oceans. It contains information on the source and effects of each tsunami.

♦ The Earth Resources Observation Systems Data Center (http://edcwww.cr.usgs.gov/) houses the National Satellite Land Remote Sensing Data Archive, a comprehensive, permanent record of the planet's land surface derived from almost 40 years of satellite remote sensing. By 2005, the total holdings will come to some 2.4 million gigabytes of data.

per form but is also available electronically. Examples include *Science On-line, Cell, Nature,* and many others.

- ♦ Electronic print (e-print) servers—preprint or reprint servers on which authors in specific fields post their articles. The original and most widely copied preprint server is the Los Alamos physics preprint server (http://arxiv.org/). Started in 1991 by Los Alamos physicist Paul Ginsparg as a service to physicists in a small subfield of physics, this server has grown to cover many fields of physics, astronomy, mathematics, and computation. Other preprint servers have been developed to serve other fields, but most fields do not use preprint servers as extensively as physics.
- ♦ Non-peer-reviewed publications on-line—includes electronic newsletters, magazines, and working papers.
- Personal Web pages—maintained by individuals or research groups. Many scholars post their own work on these sites, which may include "reprints" of published material, preprints, working papers, talks and other unpublished material, bibliographies, data sets, course material, and other information of use to other scholars.

In addition, a number of services facilitate searching and provide abstracts and (in some cases) full text of articles in paper or p-e journals. These services include LexisNexisTM, databases of journals sold to academic libraries, and public sources such as PubMed Central (<http://www.pubmedcentral.nih.gov/>) and PubSCIENCE (<http://www.osti.gov/pubsci>).9 These can be considered elements of digital libraries. See sidebar, "Digital Libraries."

An example of rapid expansion in electronic scholarly communication is the Los Alamos preprint server (http://arxiv.org/), which continues to grow in terms of both submissions and connections. As of April 2001, it was receiving more than 2,500 new submissions each month and averaging more than 100,000 connections (for searching, reading, or downloading papers) each day. It has become the main mode of communication in some fields of physics, and 17 mirror sites have been established around the world to provide alternative access to the information in it.

Kling and McKim (2000) note that one should not expect the preprint server mode of electronic communication to expand to all fields, however. High-energy physics had a culture of wide sharing of preprints before the advent of the World Wide Web, and researchers in this field now use electronic communication extensively. Molecular biologists, by contrast, traditionally shared preprints only among smaller groups and continue to rely more on paper journals. Different fields have different attitudes about posting material on the Web prior to publication in a peer-reviewed journal. In physics, such posting is standard practice; in medicine, it is viewed as dangerous because the public may make medical decisions based on non-peer-reviewed science. The absence of e-print servers in fields such as atmospheric research, oceanography, and cli-

mate science is evidence of substantial differences in scholarly communication across fields of science.

Electronic journals also have been expanding rapidly. The Association of Research Libraries 2000 directory of scholarly electronic journals and academic discussion lists (Mogge and Budka 2000) identifies 3,915 peer-reviewed electronic journals, up from 1,049 in 1997. Friedlander and Bessette (2001) cite estimates ranging from 3,200 to 4,000 e-journals in science, technology, and medicine. Most of these are not electronic-only journals but rather are electronic versions of, or supplements to, print journals.

Electronic Scholarly Communication Benefits. Electronic scholarly communication has many potential benefits. Electronic search tools make it possible for scholars to find information more easily and quickly, and scholars do not have to worry about whether journals are missing from the library. Electronic documents potentially offer richer information than print documents. They are not constrained by page limits and can contain multimedia presentations or computer programs and data as well as text, thus enriching the information and facilitating further work with it. Additional references, com-

Digital Libraries

The digital library is a concept related to both scholarly communication and scientific databases. The concept encompasses a variety of digital collections of information, including digital versions of traditional library, museum, and archive holdings. The World Wide Web is considered an extensive but rudimentary digital library because search methods typically cover only a small part of the collections (President's Information Technology Advisory Committee 2001). Newspaper archives and genomic databases also are considered digital libraries.

A vision set forth for digital libraries is that they will let all citizens, anywhere and anytime, use any Internet-connected digital devices to search and access all of human knowledge (President's Information Technology Advisory Committee 2001). Key issues in achieving this vision include:

- improving the ability to search for information, in part by improving the "metadata" systems for describing and organizing collections;
- improving the human interfaces with the libraries;
- improving the ability to store and retrieve materials across diverse independent collections;
- developing long-term storage technologies and efficient procedures for transferring ephemeral content into long-term storage; and
- determining how to manage intellectual property rights for digital collections.

⁹PubSCIENCE is a World Wide Web service developed by DOE to facilitate searching and accessing of peer-reviewed journal literature in the physical sciences and other energy-related disciplines.

ments from other readers, or communication with the author can be linked to the document.

Electronic communication is generally thought to speed the dissemination of scientific information, and this is generally thought to increase scientific productivity. However, some scientists suggest that the Web, by speeding electronic communication, can encourage scientists to rush to become part of the latest trend, leading them to abandon other paths of research too quickly (Glanz 2001b).

There are also potential advantages for libraries. Many patrons can access the same electronic information at the same time without needing to visit the library, electronic archives eliminate the space requirements of old journal collections, and electronic media help libraries stretch limited financial resources, especially for accessions.

Electronic documents also have potential economic benefits. Once a document is prepared in electronic form, the marginal cost of providing it to additional readers is very low. Electronic documents also offer the benefit of accessibility. Electronic documents can be made available over the Internet to scholars around the world who do not have access to major research libraries. For example, the Los Alamos archive is allowing scientists in geographically isolated and small institutions to participate in leading-edge research discussions (Glanz 2001a). Several publishers have announced that they will provide free electronic medical journal access to medical schools, research laboratories, and government health departments in poor countries (Brown 2001).

Electronic Scholarly Communication Issues. All of the factors mentioned above combine to exert strong pressures for making scholarly information available electronically. Although these potential benefits support the rapid expansion of electronic communication, several issues remain to be resolved, including issues related to function, economics, and archiving.

Function. Although nonrefereed electronic publications (such as preprint servers) can be much less expensive than print journals (Odlyzko 1997), such publications do not perform all of the functions of the traditional system of printed academic journals. For example, journals organize articles by field and manage peer-review processes that help to screen out bad data and research, scholars achieve recognition through publication in prestigious journals, and universities base hiring and promotion decisions on publication records. For this reason, preprint servers are not likely to replace peer-reviewed journals.

Economics. For peer-reviewed journals (in either paper or electronic form), editing and refereeing of manuscripts and general administration account for a large share of costs (Getz 1997). At least initially, these costs remain about the same for electronic journals. In addition, electronic journals have costs associated with acquiring and implementing new technology and formatting manuscripts for electronic publication.

Electronic publication also can affect the revenue stream of print publishers. If a publisher provides a site license for a university library that enables anyone on campus to read the journal, individual subscriptions from that campus may decline. Moreover, advertisers may find electronic journals less attractive than print versions.

Publishers are currently experimenting with different ways of pricing electronic journals. Some publishers provide separate subscriptions for electronic and print versions, and the price of the electronic subscription may be higher or lower than the price of the print subscription. Others provide the electronic version at no charge with a subscription to the print version. Some publishers offer free on-line access to selected articles from the print version and regard the on-line version as advertising for the print version (Machovec 1997). Publishers of feebased electronic journals generally protect their information from unauthorized access by making the journals accessible only to certain Internet domains (such as those of universities that have acquired a site license) or by using passwords.

Electronic resources represent an increasing share of library costs. The Association of Research Libraries (Kyrillidou 2000) reported that electronic resources (e.g., indexes and subscriptions to on-line journals) increased from 3.6 percent of library material expenses in 1992–93 to 10.5 percent in 1998–99. Overall, serial costs (including both paper and electronic serials) increased over this same period, from a median of \$161 per serial in 1992 to \$284 in 2000. Library budgets are under increasing pressures as they seek to satisfy demands for both paper and electronic journals.

Archiving. Another key issue is the archiving of electronic publications (Friedlander and Bessette 2001). One fundamental issue is the technical question of how to maintain records over the long term, because the electronic medium degrades and electronic formats change. Another fundamental issue is the underlying tension in electronic media between the opportunity to revise and update papers to maintain currency and the need to maintain the record. Another question to be addressed is whether an entire issue of an on-line magazine or newspaper should be preserved or whether it suffices to create a database of individual stories that can be individually retrieved but can never be reconstituted into the actual issue as it existed on the day readers first read the news. Other questions relate to responsibility for long-term preservation (whether publishers or libraries should be primarily responsible), copyright (how to issue and enforce copyrights), and maintenance (as technologies evolve, the particular technology required to view a given file may become obsolete, effectively eliminating the record).

Collaboration

Computer networking was developed as a tool for scientists and engineers, and e-mail and file transfers have long supported collaboration among scientists and engineers. Shared databases, intranets, and extranets have helped geographically separated scientists and engineers work together.

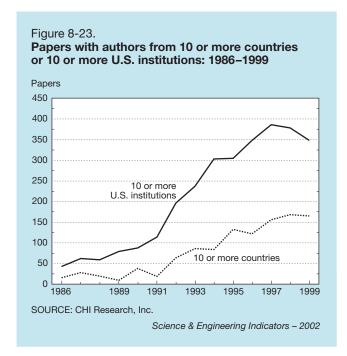
Scientific collaboration, as measured by the increase in the percentage of papers with multiple authors, has been increasing steadily for decades. (See chapter 6, "Industry, Technology, and the Global Marketplace.") Walsh and Maloney (2001) have found that computer network use is associated with more geographically dispersed collaborations as well as more productive collaborations.

Collaborations have been growing larger in a number of fields, often because scientists are pursuing increasingly complex problems and, in some cases, also because agency funding programs encourage multi-investigator and multidisciplinary research teams. These collaborations are facilitated by IT, especially e-mail and the World Wide Web. Large-scale scientific collaborations may especially benefit from new IT. The number of research papers with authors from multiple countries or institutions has increased rapidly, a trend that has coincided with the rapid expansion of the Internet. (See figure 8-23.)

Over the past decade, advanced tools have emerged to support "collaboratories"—geographically separate research units functioning as a single laboratory (CSTB 1993). These technologies allow:

- remote access to scientific instruments over the Internet, making it possible for researchers from different sites to use a single major scientific instrument (such as a synchrotron at a national laboratory) as a network of instruments operating at different places:
- ◆ Internet-based desktop videoconferencing;
- shared access to databases and computer simulation;
- shared virtual workspaces, such as "white boards" on which researchers can sketch ideas; and
- shared electronic laboratory notebooks to capture the details of experiments.

These tools were originally developed and demonstrated through several collaboratory pilot projects, including the



NSF-sponsored Space Physics and Aeronomy Research Collaboratory (http://intel.si.umich.edu/sparc/) and the DOE-sponsored Materials MicroCharacterization Collaboratory (http://tpm.amc.anl.gov/MMC) and Diesel Combustion Collaboratory (http://www-collab.ca.sandia.gov/snl-dcc.html).

The collaboratory concept has moved beyond pilot projects to the point where many new large-scale projects have collaboratory components. Many of the tools used in the early pilot projects, such as Internet-based videoconferencing, are now available in inexpensive commercial software. Examples of new major research projects that have a collaboratory component include the following:

- ◆ The NIH-funded Great Lakes Regional Center for AIDS Research, a collaboratory of Northwestern University, University of Wisconsin-Madison, University of Michigan-Ann Arbor, and University of Minnesota-Minneapolis investigators (http://www.greatlakescfar.org/cfar/).
- NIH's Human Brain Project, a cooperative effort among neuroscientists and information scientists to develop tools for brain research (http://www.nimh.nih.gov/neuroinformatics/index.cfm). This project emphasizes tools to aid collaboration between geographically distinct sites.
- ♦ The NSF-funded George E. Brown, Jr., Network for Earthquake Engineering Simulation (NEES), a national networked collaboratory of geographically distributed, shareduse experimental research equipment sites (with teleobservation and teleoperation capabilities) for earthquake engineering research and education. When operational in 2004, NEES will provide a network of approximately 20 equipment sites (shake tables, centrifuges, tsunami wave basins, large-scale laboratory experimentation systems, and field experimentation and monitoring installations) (NSF 2001b).
- ♦ The NSF-funded Distributed Terascale Facility (DTF) will be a multi-site supercomputing system. It will perform 11.6-trillion calculations per second and store more than 450-trillion bytes of data, and will to link computers, visualization systems and data at the National Center for Supercomputing Applications in Illinois, the San Diego Supercomputer Center (SDSC) in California, Argonne National Laboratory in suburban Chicago and the California Institute of Technology in Pasadena (NSF 2001c).

Although collaborative research projects are being designed around IT, it is unclear whether virtual collaborations will be as successful as colocated collaborations. Teasley and Wolinsky (2001) note that collaboratories have limits. Social and practical acceptability are the primary challenges. Collaboratories do not replace the richness of face-to-face interaction, and concerns about trust, motivation, data access, ownership, and attribution can affect collaboratory performance.

Finholt (2001) notes that, although studies of early collaboratories suggest that e-mail and computer-mediated communication enhance scientific productivity and support

larger and more dispersed collaborations, electronic communication alone is not enough to enable broader collaborations. Collaboratory technologies have not dispersed to scientific users as fast as other Internet technologies (such as e-mail or the World Wide Web), which suggests that major challenges may be involved in supporting complex group work in virtual settings. Most practices and routines of research groups assume a shared space, and transferring these practices to virtual spaces can be difficult. Collaboratories may benefit graduate students and "nonelite" scientists the most, because they are the members of the scientific community least able to afford the costs of travel. Also, the increase in outside participation that results from virtual collaboration may create distractions for top researchers.

Olson and Olson (2001) note that distance collaborations work best when the work groups have much in common, the work is loosely coupled, and the groups have laid both the social and technical groundwork for the collaboration. Lacking these elements, distance collaborations are much less likely to succeed.

Collaboratory technologies raise interesting questions about the effects of IT on the organization of science and technology (S&T). Will multi-institution, electronically enabled collaborations become the norm for large-scale science projects? Will collaboratories make science more open to nonelite scientists? How do collaboratory technologies affect the productivity of S&T?

IT and Innovation

In addition to its interactions with R&D, IT influences several other elements of the innovation process, including the market environment for and the organization of innovation. The Council of Economic Advisers (2001) notes that the U.S. economy in the late 1990s was characterized by the high rate of technological innovation and by the central role of IT. The council observes that innovation in the "new" economy appears to have changed in several ways, including the intense competition and positive feedback that drive innovation, the mechanisms for financing innovation, the sources of R&D, and the innovation process itself. IT is involved in each of these changes, and many of the changes are most visible in the IT sector.

Market Environment for Innovation

The rapid pace of technological advances, together with the expectation that this pace will continue (see sidebar, "Moore's Law"), has led to an environment in which companies in most industries know they must continually innovate. As noted above, intense competition and feedback drive the development and adoption of new technologies. The availability of one technology stimulates demand for complementary technologies, which in turn lowers production costs and encourages further demand for the initial technology.

The Internet may be stimulating innovation by forcing many industries to innovate. For example, in the food industry, the fact that some companies are using electronic procurement is forcing others to do the same (Hollingsworth 1999). In some cases, IT may increase competition simply by making markets more global and bringing firms in contact with more competitors.

Lewis (2000) notes that telecommunications and IT have accelerated business processes. Technology adoption and diffusion rates are faster than they were in previous decades. In addition, the information economy has led to network effects (see sidebar, "Metcalfe's Law") in many areas, giving a major advantage to the company that is the first to bring a new product to market. If a company is not the first to market, then it needs to match and improve on the new product very quickly. The consequence of this environment is that technology transfer must occur faster and faster. Lewis argues that corporate R&D must change its traditional way of doing business, which is too slow.

The rapid improvement in IT has created opportunities in new applications such as secure Web servers or e-commerce software, which in turn create opportunities for new businesses. New forms of business activity (such as electronic marketplaces) and new IT-enabled business processes present many opportunities for innovation.

Organization of Innovation

Dewett and Jones (2001) review the literature on how IT affects organizational characteristics and outcomes. They note that although the literature contains very little information on the specific role of IT in promoting innovation, it is possible to identify many innovation-related effects of IT on organizations, including the following:

- ◆ IT can enhance the knowledge base available to each employee, enable faster scanning and monitoring of the external environment, and improve both the employees' and the organization's knowledge of best practices and relevant leading-edge technologies.
- ♦ IT can mitigate the tendency toward specialization (which can reduce people's ability to understand the context of the organization) and also can help promote innovation by better connecting specialists to the market.
- IT may increase absorptive capacity, which is the ability of an organization to recognize the value of external information, assimilate it, and apply it commercially.
- ♦ By helping organizations codify their knowledge bases, IT can promote the diffusion of knowledge.
- ◆ IT has helped organizations streamline product design by replacing traditional sequential processes with parallel processes in which employees in different functions work simultaneously, with continual interaction through electronic communication.
- ◆ IT is changing organizational forms and allowing virtual organizations. New IT-enabled organizational forms can be more responsive to pressures such as heightened market volatility, the globalization of business, increased un-

certainty in the economy, and demographic changes in labor and consumer sectors.

In contrast, electronic communication may hinder innovation by decreasing informal communication and may also lead to information overload.

Thus, IT has many possible effects on organizations, and these effects suggest a considerable positive influence on innovation. It is important to keep in mind, however, that scholarly literature on this subject is sparse.

Johannessen, Olaisen, and Olsen (2001) suggest that because IT more effectively transfers explicit knowledge than tacit knowledge, it may lead to the mismanagement of innovation. Explicit knowledge is relatively easy to express in symbols, digitize, and transfer. Tacit knowledge is rooted in practice and experience and typically is transmitted through training and doing. Companies typically focus IT investment on the explicit portion of their knowledge base and deemphasize the tacit portion. Yet much of the research literature argues that tacit knowledge is critical in determining how well a company can innovate and compete.

IT also has led to changes in the organization of innovation beyond the boundaries of individual organizations. The Council of Economic Advisers (2001) notes that innovation traditionally has been isolated within large companies. Today, innovation increasingly is performed by both large and small companies that collaborate with each other and with academic institutions and government agencies.

With the expansion of the world's supply of scientists, technologists, and knowledge workers and of the knowledge bases available to them, access to external knowledge sources is becoming an increasingly important factor in the ability of organizations to participate in innovation. IT has helped organizations coordinate highly dispersed innovation activities by providing them with new management techniques, software, and communication systems. One aspect of the trend toward dispersion of innovation activities is the outsourcing of innovation. Pharmaceutical companies have long outsourced basic research to universities, institutes, and government laboratories. Many large pharmaceutical companies rely on small technology companies for innovation and then acquire these companies. In the computer and automotive industries, manufacturers have long relied on component makers for design and engineering work. Much of the innovation in these industries takes place at the interface between manufacturers and their innovative suppliers. IT has made outsourcing more attractive for companies (Quinn 2000) by facilitating the process with advances in modeling and simulation, collaborative tools, and management software.

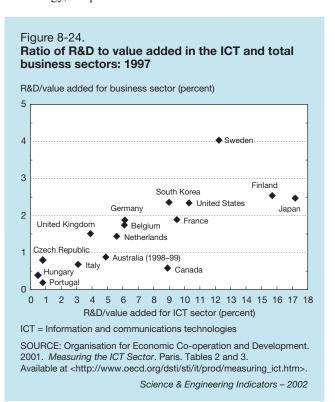
One example of new organization in innovation is open source software development (Lerner and Tirole 2001). In open source software development, the source code is made broadly available. Users can modify the software, but their modifications are also returned to the community or organization that oversees the development of the software. A number of open source software programs are widely used, including Linux (a PC operating system), Apache (Web server software), and

Sendmail (which underlies e-mail routing on the Internet). Participation in open source projects is voluntary. Although participants appear to be motivated by altruism, they do benefit from their efforts. Programmers who donate their time benefit from recognition, and companies that support the programmers benefit from improved programs and better monitoring and absorption of external technology (Lerner and Tirole 2001). A number of companies make money not by selling the software, which is freely available, but by selling complementary services (e.g., documentation, installation software, and utilities). The President's Information Technology Advisory Committee recommended that the Federal Government support open source software development for high-end computing.

Innovation in IT

The IT sector accounts for a large and growing part of R&D and innovation in the United States and other countries. The information and communication technology (ICT) sector is more R&D intensive than industry as a whole. Figure 8-24 compares the ratio of R&D to the value added for the ICT sector with the same ratio for the overall business sector in OECD countries. For most countries, the ICT sector is about five times more R&D intensive than the business sector as a whole; however, countries vary widely in the R&D intensity of their IT industries. Some of the countries that are the most innovative in IT, including Sweden, Finland, Japan, and the United States, have the most R&D-intensive ICT industries.

Analyses of patent data suggest that innovation in IT is somewhat different from innovation in other areas of S&T. Hicks et al. (2001) found that compared with other areas of technology, IT patents cite scientific literature less exten-



sively.¹⁰ In addition, the median age of patents cited by IT patents (termed the technology cycle time) is less than the median age of scientific papers cited (termed the science cycle time). Technology cycle times are faster in IT (6–6.5 years) than in other areas of technology. The analysts concluded that IT patents cite other technology patents more extensively than scientific papers because IT is moving too fast for scientific research to keep up.

Growth in IT patenting activity does not seem to be accompanied by growth in research publishing activity (Hicks et al. 2001). Based on documents referenced in patents, IT patents seem to draw on a particularly diverse set of nonpatent, nonresearch technical documentation that includes nonpatented software. It appears that nonresearch technical work may underlie innovation in IT more extensively than is the case in other technologies and that IT innovation is less directly dependent on scientific research than are many other technologies. Hicks et al. (2001) also note that patenting (one measure of innovative activity) is accelerating in IT. The IT patents share of all U.S. patents increased from 9 percent in 1980 to 25 percent in 1999. IT patents per \$1 million of company R&D expenditures nearly doubled between 1990 and 2000. Similar increases were not observed in other areas. Although such statistics might simply indicate an increased propensity to patent in the IT sector, the extent to which IT patents are cited by other patents has increased, which suggests that the quality of IT patents has not deteriorated.

IT and Higher Education

IT pervades higher education. As the demand for IT workers has grown, university priorities have shifted accordingly, and a separate certification and training system for IT workers also has emerged. IT is increasingly used in instruction, and distance education continues to expand. IT may lead to further restructuring of colleges and universities. This section highlights some of the ways in which IT is affecting higher education.

IT Credentialing

Adelman (2000) analyzes the new system of credentialing that has arisen in ICT industries during the past decade. Companies and industry or professional associations have created more than 300 discrete certifications since 1989. Approximately 1.6 million individuals had earned about 2.4 million IT certificates by early 2000, most since 1997. Students outside the United States earned about half of the certificates. To earn a certificate, a candidate must pass an exam administered by a third party. A large industry has arisen to prepare candidates for these exams. This industry includes organizations that provide courses, tutorials, practice exams, self-study books, and CD-ROMs. Although some traditional four-year colleges and community colleges prepare students for these certification exams, much of the industry that supports IT certification is outside higher education as traditionally defined.

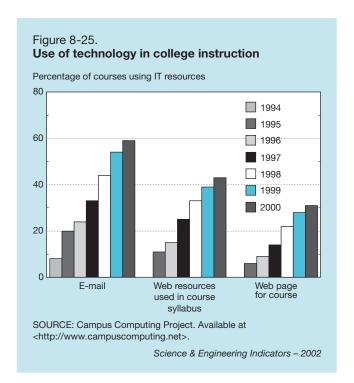
IT in Instruction

The Campus Computing Project (2000) found that IT use in college courses is increasing. There are indications that IT use is leveling off; nevertheless, e-mail, the Internet, and course Web pages are being used in more courses every year. (See figure 8-25.)

In some cases, decisions about IT use are left to individual professors. However, some universities (such as the University of California, Los Angeles) have required professors to establish Web pages for each course and to put syllabuses online. Support for the increased use of IT on college campuses has not been universal. Many professors and administrators enthusiastically embrace new technologies, and others prefer to wait for other institutions to find out which new technologies are useful in improving the quality of education.

Much of the new IT being used in scholarly communication and research can be used in instruction as well. Students can use on-line scholarly literature, participate in on-line scientific experiments, and learn from computer modeling and simulation. The future is likely to bring many additional IT applications in instruction.

Kulik (forthcoming) reviewed 44 studies from the 1990s on the effects of instructional technology in college courses. The studies focused on five computer applications: computer algebra systems, computer use (tutoring, simulations, and animations) in science, and computer-assisted language learning. In each study, instructional outcomes for students taught with and without computer help were compared. The instructional outcome measured most often in the studies was student learning. Kulik found that over the years, instructional technology has proven to be more and more effective in improving learning in college courses. Studies in the 1990s show a greater positive effect of instructional technology than stud-



¹⁰IT in this discussion of patents consists of computers, peripherals, telecommunications, semiconductors, electronics, and software. Hicks et al.'s analysis covers patent activity between 1980 and 1999.

ies in the 1980s and earlier decades. The growing effectiveness of instructional technology coincides with the dramatic improvements in computing and with improvements in instructional software.

Although computer technologies are helping to improve student learning, they can also make it easier for students to cheat and plagiarize. The Internet contains many collections of school papers that students can download and use for their classes.

Distance Education

Distance education is not new. An estimated 100 million Americans have engaged in distance study, mostly correspondence courses, since 1890 (Distance Education and Training Council 1999), and in the 1960s there was widespread optimism about the use of television in education. IT is providing significant new tools for distance education. Many schools are either establishing distance education programs for the first time or expanding existing programs.

In on-line distance courses, the instructor typically e-mails "lectures" or posts them on a website, and students submit assignments and have "discussions" via e-mail. Courses often supplement textbooks with Web-based readings. Participants also may meet in a chat room at a certain time for on-line discussions. Courses also may have on-line bulletin boards or Web conferences, in which participants ask and respond to questions over time. In the not-too-distant future, as Internet bandwidths increase, video lectures and videoconferencing will become more common in on-line courses. Some courses may use more elaborate systems (so-called MUD/MOOs¹¹) for group interaction, as well as groupware programs that involve simultaneous viewing of graphics and use of a shared writing space (e.g., white boards) (Kearsley 2000). Some courses may also use computer simulations over the Internet.

Distance education offers several potential advantages: it allows students to take courses that are not available locally; it allows students to balance coursework with their career and family life; and it can make education more available to people who are employed, especially those who are older and in midcareer or those who have family responsibilities. For universities, it offers a way to expand enrollment without increasing the size of the physical plant.

Although distance education traditionally is regarded as involving the delivery of courses to remote locations, the techniques of distance education, especially on-line education, can be incorporated in on-campus instruction as well. Universities are finding that significant numbers of on-campus students sign up for distance education courses when they are offered. At the University of Colorado in Denver, for example, more than 500 of 609 students enrolled in distance education courses were also enrolled in regular on-campus courses (Guernsey 1998). On-line courses can be more convenient for on-campus students, giving them greater flexibility in scheduling their time. Professors can augment their on-line courses with Web-based materials and guest lecturers in remote sites.

Distance Education Trends. The National Center for Education Statistics has conducted two surveys of distance education in postsecondary education institutions: the first in the fall of 1995 and the second in the 1997/98 academic year (National Center for Education Statistics (NCES) 1999b). The first survey covered only higher education institutions, but the second survey covered all postsecondary educational institutions. These surveys document that distance education is now a common feature of many higher education institutions, and its popularity is growing. The majority of courses are at the undergraduate level and are broadly distributed across academic subjects.

The number of higher education institutions offering distance education is growing. In 1997/98, 44 percent of all two-and four-year institutions offered distance education courses compared with 33 percent in fall 1995. Distance education is more widely used in public four-year institutions than in private four-year institutions, but private institutions are also increasing their use of it. In 1997/98, distance education was offered by 79 percent of public institutions (compared with 62 percent in fall 1995) and 22 percent of private institutions (compared with 12 percent in fall 1995).

Distance education course offerings and enrollments are growing more rapidly than the number of institutions that offer distance education. The number of courses offered in two-and four-year higher education institutions doubled from 25,730 in fall 1995 to 52,270 in 1997/98. The increases were fairly similar across all categories of institutions (two- and four-year, public and private, and all enrollment-size categories). Course enrollments also increased sharply, more than doubling from 753,640 in fall 1995 to 1,632,350 in 1997/98 (NCES 1999b).

The availability of degrees that can be completed exclusively with distance education courses has remained essentially constant. Of higher education institutions that offer distance education, 23 percent offered degrees in fall 1995 and 22 percent did so in 1997/98 (NCES 1999b).

Technologies used for distance education have changed significantly. In fall 1995, the most widely used technologies were two-way interactive video (57 percent) and one-way prerecorded video (52 percent). These were still widely used in 1997/98 (56 and 48 percent, respectively). Internet-based courses, however, expanded greatly. Of all the institutions that offered distance education courses in 1997/98, 60 percent offered asynchronous (not requiring student participation at a set time) computer-based instruction and 19 percent offered synchronous (real-time) computer-based instruction (NCES 1999b).

Significance of On-line Distance Education. Despite substantial (and growing) experience with on-line distance education, thorough assessments of its effectiveness have been relatively few. Existing evidence suggests that, at least in some circumstances, it can be very effective. The rapid growth and reported success of some on-line distance education programs indicate that they are providing acceptable learning experiences. A review of the literature on on-line classes (Kearsley, Lynch, and Wizer 1995) found that compared with traditional classes, student satisfaction was higher, measured student achievement

¹¹MUD stands for multi-user domain or multi-user dungeon (reflecting its origins in games), and MOO stands for MUD, object-oriented.

was the same or better, and student-instructor discussions usually were more frequent. On the other hand, some case studies document that on-line distance education can be frustrating for both students and instructors. The growth of on-line distance education has far-reaching implications for higher education. Although on-line education may expand the pool of people who have access to education, it may also take students away from traditional education. Some scholars express concern that it will undermine the traditional college experience. Some question whether it can match the quality of face-to-face instruction. Moreover, the kind of intellectual and social community that characterizes the college experience may be much harder to achieve through distance learning.

IT Issues for Universities

IT in general and distance education in particular raise new issues for universities. Distance education brings universities into competition with each other in a new way. Because distance education courses are available to anyone anywhere, they allow universities to compete for students outside their own geographic areas. Top-tier universities such as Stanford and Duke are marketing Internet-based master's degrees to national audiences. New distance education-based universities such as Jones International University (http:// www.jonesinternational.edu>), the first on-line-only university to gain accreditation; the University of Phoenix on-line (<<http://online.uophx.edu>>); and Western Governors University (<http://www.wgu.edu>) are marketing courses that compete with the continuing education services of universities and colleges that in the past had been the only providers of such services in their regions. Some distance education providers see opportunities to market American university degrees to large student populations abroad. The reverse is also happening: the United Kingdom's Open University, which began providing distance education in the United Kingdom in 1971 and has established a good reputation there, has started an operation in the United States (Blumenstyk 1999a). In contrast to many institutions that are viewing Web-based course materials as a new source of revenue, MIT announced in 2001 that it would make nearly all of its course materials available for free on the Web over the next ten years (Massachusetts Institute of Technology, 2001).

In addition, distance education is creating new markets for companies that sell print materials and software to assist in on-line courses (Blumenstyk 1999b). Publishers such as McGraw-Hill and software companies such as Microsoft and Oracle have developed and are marketing on-line courses (Morris 1999). These commercial on-line courses represent another potential source of competition for universities, especially in preparing students for IT credentialing.

Distance education technologies also raise questions about the role of professors. Some view these technologies as new tools for professors. Others, however, foresee "mass production" education in which packaged multimedia courses will reduce the importance of professors (Noble 1998). The expanding and potentially lucrative new market for on-line course materials raises the issue of whether professors or the university should own the intellectual property embodied in on-line courses. The American Association of University Professors has taken the position that professors rather than institutions should retain primary property rights for on-line course materials (Schneider 1999) and has questioned the accreditation of Jones International University (Olson 1999).

Brown and Duguid (2000) note that colleges and universities provide three essential functions to learners: access to an authentic community of learning, resources to help learners work within these communities, and widely accepted representations of learning and work (such as degrees and transcripts). Brown and Duguid also note that many proposals for new "virtual universities" fail to provide one or more of these functions. Conventional universities serve all of these functions by combining five elements: students, faculty, research, facilities, and an institution able to provide an accepted degree. Brown and Duguid suggest that these elements will remain but that new technologies will allow the elements to be in a looser configuration, not necessarily combined in a single collocated organization.

The IT Workforce

During the 1996–2000 period, the rapid expansion of IT development and application during a period of full employment in the overall economy led to concerns about the availability of IT workers. In 2001, however, the cooling of the economy (especially in the IT sector) has, at least temporarily, ameliorated these concerns.

The Bureau of Labor Statistics has projected the future demand for IT workers (U.S. DOC 1997, 1999b, 2000c) for six core occupational classifications: computer engineers, systems analysts, computer programmers, database administrators, computer specialists, and all other computer scientists. These projections indicate that between 1998 and 2008, the United States will require more than 2 million new workers in these six occupations.

One indicator of the supply of IT workers is the number of computer science degrees awarded. After increasing sharply in the early 1980s, that number declined sharply after 1986 and has only begun to increase again since 1996. (See chapter 2, "Higher Education in Science and Engineering.")

The IT industry asserted that a serious shortage of IT workers exists, and many companies in various industries indicated that they needed more IT-trained workers to meet the growing demand. However, the existence of a shortage of IT workers was the subject of debate. Some employee groups believed there were enough trained technical professionals but that industry had not tapped existing labor pools (especially older engineers). The debate has been especially polarized over the issue of whether to allow more foreign workers with technical training to enter the country on temporary H-1B visas.

Several studies have examined the IT workforce issue (CSTB 2001; Freeman and Aspray 1999; Johnson and Bobo 1998; Lerman 1998; U.S. DOC 1999b). (See also chapter 3,

"Science and Engineering Workforce.") These studies generally reached the following conclusions:

- ◆ During 1996–2000, the IT labor market was somewhat tighter than the overall labor market. Existing data, however, cannot prove or disprove that such a shortage existed. Federal data are limited by untimely reporting, outof-date occupational descriptions, and incompatibilities in supply-and-demand data collected by different agencies.
- ◆ The IT labor market is not homogeneous. Supply-and-demand characteristics vary by region, industry segment, and specific skill. Because IT product cycle times are very fast, the industry pays a premium for people who already have specific current skills and do not require training to be effective. Competition is especially intense for people with specific "hot" skills in specific markets.
- ♦ People enter IT careers in a variety of ways. IT workers include people who majored in IT-related disciplines at the associate, bachelor's, master's, and doctoral degree levels; people from other science, engineering, and business fields; and people from nontechnical disciplines who have taken some courses in IT subjects. Many IT workers enter the field through continuing education programs and forprofit schools. Workers are taking advantage of new modes of instruction delivery such as distance learning.

Labor markets tend to be cyclical. In response to the tight conditions in the IT labor market during 1996–2000, wage increases attracted more people to the field, and many initiatives around the country were set up to help expand the IT workforce. Slower growth and even layoffs in the IT industry have also reduced demand for IT workers.

Conclusion

IT continues to develop rapidly as the key underlying technologies of semiconductors, disk drives, and network communications improve at exponential rates. Constant improvements in the underlying technologies make possible new IT applications that affect all areas of society, including the economy, households, government, and the R&D enterprise.

Throughout society, the utility of IT applications tends to advance much more slowly than the underlying technologies. A doubling of processing speeds, for example, does not bring a doubling of utility. The effective implementation and use of IT are the result of a complex process that requires not only adoption of a technology but also changes in organizations and institutions. As part of this process, individuals and organizations actively adapt (and sometimes resist) the technologies. As a result, the effects of IT on society often take place more slowly than visionaries predict. Nevertheless, the effects—driven by the continual change in underlying technologies—are substantial over time.

Selected Bibliography

- Adelman, C. 2000. A Parallel Postsecondary Universe: The Certification System in Information Technology. Washington, DC: U.S. Department of Education.
- Alberts, D.S., and D.S. Papp, eds. 1997. *The Information Age: An Anthology on Its Impact and Consequences*. Washington, DC: National Defense University.
- Allen, T., and M.S. Morton, eds. 1994. *Information Technology and the Corporation of the 1990s*. New York: Oxford University Press.
- Alpar, P., and M. Kim. 1991. "A Microeconomic Approach to the Measurement of Information Technology Value." *Journal of Management Information Systems* 7(2): 55–69.
- Ante, S.E., with A. Borrus and R.D. Hof. 2001. "In Search of the Net's Next Big Thing." *Business Week* (March 26): 141.
- Atkinson, R.D., and R.H. Court. 1998. *The New Economy Index: Understanding America's Economic Transformation*. Washington, DC: Progressive Policy Institute.
- Attewell, P., and J. Rule. 1994. "Computing and Organizations: What We Know and What We Don't Know." In C. Huff and T. Finholt, eds., *Social Issues in Computing*. New York: McGraw-Hill.
- Autor, D., L. Katz, and A. Krueger. 1997. "Computing Inequality: Have Computers Changed the Labor Market?" NBER Working Paper No. 5956. Cambridge, MA: National Bureau of Economic Research.
- Baily, M.N., and A.K. Chakrabarti. 1988. *Innovation and the Productivity Crisis*. Washington, DC: Brookings Institution Press.
- Baker, E.L., M. Gearhart, and J.L. Herman. 1994. "Evaluating the Apple Classrooms of Tomorrow." In E.L. Baker and H.F. O'Neil, Jr., eds., *Technology Assessment in Education and Training*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Banker, R.D., R.J. Kauffman, and M.A. Mahmood. 1993. *Strategic Information Technology Management: Perspectives on Organizational Growth and Competitive Advantage*. Harrisburg, PA: Idea Group Publishing.
- Berman, E., J. Bound, and Z. Griliches. 1994. "Changes in the Demand for Skilled Labor Within U.S. Manufacturing: Evidence From the Annual Survey of Manufacturers." *Quarterly Journal of Economics* 109(May): 367–98.
- Berman, H.M., J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, and P.E. Bourne. 2000. "The Protein Data Bank." *Nucleic Acids Research* 28(1): 235–42. Available at http://nar.oupjournals.org/cgi/content/full/28/1/235.
- Biemiller, L. 1999. "California State U. Adopts New Model to Pay for Journals." *Chronicle of Higher Education* July 16: A26.
- Blumenstyk, G. 1999a. "Distance Learning at the Open University." *Chronicle of Higher Education* July 23: A35.